AD-770 584

BAYESIAN DATA ANALYSIS OF GAMBLING PREFERENCES

Dirk Wendt

Michigan University

Prepared for:

Office of Naval Research
Advanced Research Projects Agency

2 November 1973

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

011313-3-T

THE UNIVERSITY OF MICHIGAN

AD 770584

ENGINEERING PSYCHOLOGY LABORATORY

Technical Report

Bayesian Data Analysis of Gambling Preferences

DIRK WENDT

Sponsored by:

Advanced Research Projects Agency
Department of Defense

Monitored by:

Engineering Psychology Programs
Office of Naval Research
Contract No. N00014-67-A-0181-0049
NR 197-021

Administered through:

November 1973

OFFICE OF RESEARCH ADMINISTRATION · ANN ARBOR

56

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 011313-3-T	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) BAYESIAN DATA ANALYSIS OF GAMBLING PREFERENCES		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER None
7. AUTHOR(s) Dirk Wendt		8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0181-0049
9. PERFORMING ORGANIZATION NAME AND ADDRESS Engineering Psychology Laboratory Institute of Science & Technology Ann Arbor, Michigan 48105		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 197-021 ARPA Order No. 2105
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia		12. REPORT DATE 2 November 1973
		13. NUMBER OF PAGES 51
14. MONITORING AGENCY NAME AND ADDRESS (if different from Controlling Office) Engineering Psychology Programs Office of Naval Research Department of the Navy Arlington, Virginia		15. SECURITY CLASS. (of this report) Unclassified
		15a DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bayes's theorem
Learning theories
Data analysis
Choice-among-gambles

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper emphasizes the use of Bayesian data analysis for experiments with choices among gambles. In an introductory example, the method is illustrated by a comparison of two learning theories. Special problems arise with the analysis of data from decision making experiments which assume deterministic choice models which cannot be handled by Bayesian analyses. Several ways around these difficulties are suggested, discussed, and demonstrated on two sets of data from choice-among-gambles experiments.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Bayesian Data Analysis of Gambling Preferences

Introduction

Bayesian data analysis has been feasible since 1763 when Rev. Thomas Bayes formulated his theorem (which is just a straightforward application of the definition of conditional probability):

$$P(H|D) \;=\; P(D|H)\;P(H) \;/\; \underbrace{\sum_i P(D|H_i)\;P(H_i)}_{=\; P(D)\ \text{(overall prob.)}}$$

Despite its availability for such a long time, research workers have made little use of it. Even most researchers who consider themselves Bayesians have used it only as a normative model for human information processing but not for processing data, although Edwards, Lindman & Savage (1963) have pointed out its advantages for statistical inference almost 10 years ago, and although easily readable textbooks are available now (e.g., Hays & Winkler 1970 have a long chapter on Bayesian inference, and the books by McGee (1971) and Winkler (1972) are especially devoted to these procedures).

Bayesian statistics differs from traditional statistics in using information not contained in the sample, namely, $P(H)$, the prior probability of the hypothesis. In testing hypotheses, traditional statisticians use only $P(D|H)$, rejecting a hypothesis H_i when $P(D|H_i)$ plus the probability of more extreme data is below a certain prefixed level α.

Traditional statisticians have occasionally objected to the idea of

taking into account any prior information, like $P(H_i)$, which was not obtained from an observed sample. Those who use Bayesian methods but insist upon priors inferred from previous observations rather than intuition call themselves Empirical Bayesians (e.g., Martiz, 1970).

In a sense, Bayesian statistics can be viewed as an extension of traditional statistics; it uses the same information plus something more, namely prior probabilities, under assumption that all information available should be used for decisions among competing hypotheses. Actually, according to the principle of stable estimation, even strongly biassed priors cannot do much harm to the posteriors as long as the data used for their revision do have enough diagnostic impact, and as long as the prior distribution is not too small in the region favored by the data, and/or not too peaked elsewhere. (For more details about the principle of stable estimation, see Edwards, Lindman & Savage, 1963.) Thus, the arbitrary and intuitive nature of prior distributions does not constitute a reason for not using Bayesian statistical methods.

It is probably easy to show that every scientist observing and analyzing data has some priors with respect to his hypotheses—however, to discuss this is not the point of this paper, and the reader interested in these problems is referred, e.g., to Kuhn (1962). Convenient techniques to elicit and assess the scientist's prior probability distributions over hypotheses are available; some of them are described, e.g., in Winkler (1967) and Staël von Holstein (1970).

In this paper, we pay little attention to prior distributions over

hypotheses. We will rather concentrate on likelihoods $P(D|H_i)$, which are more public and less controversial than prior $P(H_i)$.

Usually, a hypothesis to be tested in traditional statistics implies that a certain parameter value obtains, e.g., in traditional null hypothesis testing the hypothesis is: $H_o:\theta = \theta_o$ for some parameter θ, which is tested against the rather diffuse alternative that $\theta \neq \theta_o$. In most cases, traditional statisticians cannot figure a probability for the data observed given this diffuse alternative hypothesis, and therefore β, the probability of an error type II, is left unknown.

In such a case, the Bayesian usually would not consider a point hypothesis $\theta = \theta_o$ as opposed to a continuum of other values of θ, but rather would assess a continuous prior distribution over the whole parameter space, which is then treated as a continuous set of hypotheses. The evidence from the sample observed would then be used to revise this continuous prior distribution over the parameter space according the Bayes's theorem, which reads for the continuous case:

$$f(\theta|x) = \frac{g(x|\theta)\ f(\theta)}{\int g(x|\theta')\ f(\theta')\ d\theta'}$$

and gives a continuous posterior distribution over the same parameter space. Although Bayesian statistics can handle any number of competing hypotheses simultaneously—up to an infinite number which is the continuous case discussed just above—the most convenient case deals with only two competing hypotheses—such as the traditional test of H_o against its alternative, the catch-all hypothesis. The advantage of testing only two hypotheses against each other in

Bayesian analysis is that Bayes's theorem can then be written in ratio form so that P(D) cancels out:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(H_1)}{P(H_2)} \cdot \frac{P(D|H_1)}{P(D|H_2)}$$

This is known as the odds-likelihood-ratio form of Bayes's theorem:

$$\Omega_D = \Omega_o \cdot LR(D); \text{ in words:}$$

posterior odds = prior odds x likelihood ratio.

For conditionally independent data, the likelihood for the whole set of data $D = (d_1, d_2, \ldots, d_m)$ is the product of the likelihoods of the individual data d_j:

$$P(D|H_i) = \prod_j P(d_j|H_i),$$

and then the odds-likelihood-ratio equation becomes:

$$\Omega_D = \Omega_o \cdot \prod_j LR(d_j).$$

Bayesian data analysis with these formulae are easy, straightforward, and efficient if you have perfect knowledge of the data generating process which gives you P(D|H), but can be quite a problem if you don't.

Bayesian Analysis of Learning Data

Let's look at an easy case first: excellent examples to do Bayesian data analyses are comparisons of learning models. E.g., Restle & Greeno (1970)

compare a linear operator model (H_1) by Bower (1961) (also, see Atkinson, Bower & Corothers, 1965, p. 91).

$$P_n(c|H_1) = a - (a - b)(1 - \theta_1)^{n-1}$$

and an accumulative model (H_2)

$$P_n(c|H_2) = \frac{b + \theta_2 a(n - 1)}{1 + \theta_2(n - 1)}$$

where $P_n(c|H_i)$ is the probability of a correct response on trial n under the respective models, θ_i is a parameter of the learning curve, and a and b are initial and asymptotic success probabilities, respectively. Corresponding probabilities of wrong responses (errors) are $P_n(e|H_i) = 1 - P_n(c|H_i)$.

Bower (1961) had 29 Ss learn a list of 10 items, "to a criterion of 2 consecutive errorless cycles. A response was obtained from the S on each presentation of an item" (p. 528). Stimuli were pairs of consonant letters, responses were the integers 1 and 2, each of the assigned to 5 of the stimuli.

Twenty-nine Ss times 10 items makes 290 on each trial (unless some Ss did not get to the last trials because they completed their two errorless cycles earlier). The data Bower obtained, in terms of relative frequencies of correct responses on the n-th trial, are reproduced in Table 1, column 2, from Restle & Greeno (1970, p. 8).

To evaluate the two competing learning theories H_1 and H_2 given the evidence from these data, Restle & Greeno (1970) assumed a = 1, and b = .5, estimated θ_i from the data, and calculated $P_n(c|H_2)$ using these parameter estimates. Resulting $P_n(c|H_1)$, $P_n(c|H_2)$, and corresponding $P_n(e|H_1)$ and $P_n(e|H_2)$ are

Table 1: Bayesian analysis of Bower's data from Restle & Greeno (1970)

(1) trial #n	(2) P_n(hit) observed (relative frequency)	(3) $P_n(\text{hit}\mid H_1)$ predicted by Model 1	(4) $P_n(\text{hit}\mid H_2)$ predicted by Model 2	(5) $P_n(\text{miss}\mid H_1)$ [=1-(3)]	(6) $P_n(\text{miss}\mid H_2)$ [=1-(4)]	(7) $\log P_n(\text{hit}\mid H_1)$ [=log(3)]	(8) $\log P_n(\text{hit}\mid H_2)$ [=log(4)]	(9) $\log P_n(\text{miss}\mid H_1)$ [=log(5)]	(10) $\log P_n(\text{miss}\mid H_2)$ [=log(6)]	(11) f(hit) [=290·(2)]	(12) f(miss) [=290-(11)]	(13) $\log L_D$ [=(11)·(7) -(11)·(8) +(12)·(9) -(12)·(10)]
1	.50	.50	.50	.50	.50	-.3010	-.3010	- .3010	- .3010	145	145	0
2	.67	.67	.74	.33	.26	-.1739	-.1308	- .4815	- .5850	194	96	1.5746
3	.80	.78	.84	.22	.16	-.1079	-.0757	- .6576	- .7959	232	58	0.5510
4	.85	.85	.87	.15	.13	-.0706	-.0605	- .8239	- .8861	247	43	0.1799
5	.90	.90	.89	.10	.11	-.0458	-.0506	-1.0000	- .9586	261	29	0.0522
6	.93	.93	.91	.07	.09	-.0315	-.0410	-1.1549	-1.0458	270	20	0.3830
7	.95	.96	.92	.04	.08	-.0177	-.0362	-1.3979	-1.0969	276	14	0.8920
8	.96	.97	.93	.03	.07	-.0132	-.0315	-1.5229	-1.1549	278	12	0.6714
9	.97	.98	.94	.02	.06	-.0088	-.0269	-1.6990	-1.2218	281	9	0.7913
10	.98	.99	.95	.01	.05	-.0044	-.0223	-2.0000	-1.3010	284	6	0.8896
11	.99	.99	.95	.01	.05	-.0044	-.0223	-2.0000	-1.3010	287	3	3.0403

$\log LR = 9.0253$

reproduced in columns 3-6 of Table 1. Restle & Greeno then compared the two models by calculating the sum

$$\Delta_i = \sum_n (P_n(c|H_i) - P_n(c \text{ observed}))^2$$

for both models ($i = 1, 2$). Δ_1 was .0042, Δ_2 was .011, indicating a better fit of H_1.

A Bayesian data analysis would consist of calculating likelihood ratios $P_n(c|H_1)/P_n(c|H_2)$ for each correct response observed, and $P_n(e|H_1) / P_n(e|H_2)$ for each error response, and multiplying them all together to get the overall likelihood ratio.

To do so, we need absolute frequencies of errors and correct responses on the 11 trials, which are not given in Restle & Greeno's book, nor in Bower's paper. We reconstructed them by multiplying the relative frequencies given in Restle & Greeno (column 2 in Table 1) by 290 (29 Ss times 10 items), resulting in the absolute frequencies of correct responses of $f_n(c)$ and errors ($f_n(e)$) reproduced in columns 11 and 12 of Table 1. (These estimates may contain some errors if some Ss quit before reaching the 11th trial because they had completed their two errorless cycles earlier.)

For convenience, the calculation of $LR(d_j)$ and $LR(D)$ is performed in logarithms: In column 13, we have

$$\log LR(D_n) = f_n(c) [\log P_n(c|H_1) - \log P_n(c|H_2)] + f_n(e) [\log P_n(e|H_1) - \log P_n(e|H_2)],$$

and

$$\sum_n \log LR(D_n) = \log LR(D),$$

with the respective logarithms in columns 7 through 10, and observed frequencies $f_n(c)$ and $f_n(e)$ in columns 11 and 12.

The resulting log LR(D) is 9.0253, indicating a likelihood ratio LR(D) over a billion: $LR(D) \simeq 1.061 \cdot 10^9$. I.e., if we had assumed equal priors, $P(H_1) = P(H_2) = .5$, this would mean that H_1 is over a billion times more likely that H_2.

Although this could be taken as strong evidence for the principle of stable estimation—even very heavily biassed priors would have been corrected by such a large likelihood ratio, we have to consider it with some reservation."

As we pointed out already, it is doubtful if we can actually assume 290 observations in the last trials (7-11) because some Ss may have quit earlier. Reduction of the numbers of observations in the last trials would reduce LR(D) considerably because trials $n = 7$ through $n = 11$ contribute most to LR(D), except for $n = 2$.

Unfortunately, the original complete data are no longer available. However, a letter from Bower assures that these figures actually can be taken as numbers of correct responses assuming that the subjects would not make any more errors had they continued after their last two errorless cycles.

Another question is whether we really can assume independence of observations enabling us to multiply likelihoods. Although the observation themselves are clearly obtained independently, the independence assumption for the conditional probabilities $P_n(d_j|H_i)$ might not hold.

A way out of this might be not to calculate the whole learning curve for each model, but rather just to predict $P_{n+1}(d_j|H_i)$ from the P_n (observed so far) by

$$P_{n+1}(c|P_n,H_1) = (1-\theta_1)\,P_n + \theta_1 a, \text{ and}$$

$$P_{n+1}(c|P_n,H_2) = \frac{R_n + a\,\theta_2\,(R_1+W_1)}{(R_n + a\,\theta_2\,(R_1+W_1)) + (W_n + (1-a)\,\theta_2\,(R_1+W_1))}$$

In Model 2, this requires an additional assumption about R_1 and W_1; we used $R_1 = W_1 = 5$ for the calculation of $P_n(c|P_{n-1},H_2)$. Actually, the choice of $W_1 = R_1$ does not make much of a difference.

We use this example to demonstrate a slightly different way of performing the data analysis: In Table 1 we took logarithms of $P_n(c|P_{n-1},H_i)$ and $P_n(e|P_{n-1},H_i)$ for $i = 1, 2$, and then subtracted the logarithms of these probabilities for $i = 2$ from those for $i = 1$ (multiplied by the respective numbers of observations); in Table 2 we calculate the likelihood ratios for correct responses and errors directly (by dividing the hit probabilities in column 5, and by dividing the error probabilities in column 6 by those in column 7 to yield column 8), and then take the logarithms of these likelihood ratios for hits and errors (columns 10 and 12) to multiply them to the respective numbers of observations (columns 9 and 11), and sum over these products.

The log likelihood ratio is now "only" 2.2508, indicating a likelihood ratio of 178.2 in favor of Model 1. Of course, taking into account the observed number of correct responses on the previous trial in each calculation of

Table 2

(1) trial #n	(2) P_n(hit) observed	(3) P_n(hit| P_{n-1} obs.) predicted by Model 1	(4) P_n(hit| P_{n-1} obs.) predicted by Model 2*	(5) [=(3)|(4)] LR (hit)	(6) [=1-(3)] P_n(miss) predicted by Model 1	(7) [=1-(4)] P_n(miss) predicted by Model 2	(8) [=(6)|(7)] LR (miss)	(9) [=290·(2)] f(hit)	(10) [=log(5)] log LR(hit)	(11) [=290-(9)] f(miss)	(12) [=log(8)] log LR(miss)
1	.50							145		145	
2	.67	.67	.74	.9054	.33	.26	1.2692	194	- .0431	96	0.1035
3	.80	.78	.78	1.0000	.22	.22	1.0000	232	0	58	0
4	.85	.87	.85	1.0235	.13	.15	.8667	247	0.0103	43	- .0621
5	.90	.90	.88	1.0227	.10	.12	.8333	261	0.0098	29	- .0793
6	.93	.93	.92	1.0109	.07	.08	.8750	270	0.0048	20	- .0580
7	.95	.95	.94	1.0106	.05	.06	.8333	276	0.0045	14	- .0793
8	.96	.97	.96	1.0104	.03	.04	.7500	278	0.0043	12	- .1249
9	.97	.97	.96	1.0104	.03	.04	.7500	281	0.0043	9	- .1249
10	.98	.98	.97	1.0103	.02	.03	.6667	284	0.0043	6	- .1761
11	.99	.99	.98	1.0102	.01	.02	.5000	287	0.0043	3	- .3010

*assuming $R_1 + W_1 = 10$

$\Sigma = 2.2508 = \log LR$

$\hat{LR} = 178.2$

$P_n(c|H_i, P_{n-1})$ brings these probabilities under both models closer to the actual data, and thus levels out differences between them. The resulting likelihood ratio is still large enough to correct even strongly biassed prior odds against Model 1, and now it takes conditioned non-independence into account. The analysis could be further improved by many maximum likelihood extimates for θ_i rather than the least squares estimates we took from Restle & Greeno (1970) for this demonstration. However, since the evaluation of learning models is not our main concern in this paper, we will now turn to analyses of choice-among-gambles data.

Bayesian Analysis of Gambling Preferences

As we have seen, Bayesian data analyses are quite straightforward models that provide us explicit probabilities of occurrence between 0 and 1 for each event we might observe. We have taken learning curves as an example; other feasible examples could be taken from psychophysics, signal detection theory, Lucean & Thurstonean choice theories, etc.

However, in analyzing gambling preference data we encounter different problems, particularly with deterministic choice models. Since they require deterministic choices, i.e., with probabilities 0 and 1, no Bayesian data analysis is feasible under these assumptions. This may be one of the reasons why decision analysts and other scientists strongly advocating Bayesian procedures as normative models for human information processing rather seldom use Bayesian methods in their data analyses: they mostly favor deterministic choice models which prevent them from applying their own principles.

We are going to illustrate Bayesian data analyses of choice-among-gambles data on two sets of data here, both borrowed from colleagues: one is from an experiment by Hommers (1973) with normal and educable retarded children of 8, 10, 12, and 14 years of age where it seems rather appropriate to replace the deterministic normative model by a probabilistic one, the other set of data is from an experiment by Seghers, Fryback & Goodman (1973) with adult subjects where the conventional (Lucean) probabilistic choice models might indicate too weak preferences as compared to the choice probabilities inferred from the data.

Hommers' Data

Hommers (1973) in his dissertation compares choices among bets made by 8, 10, and 12 years old normal children, and 8, 10, 12 and 14 years old educable retarded children. Each set of gambles presented as choice alternatives to the S consisted of 3 bets labelled W, L, and S, respectively, where W indicates the choice with the largest amount to be won but with the smallest winning probability, S the one with the largest winning probability but the smallest amount, and L had medium probability and payoff. Table 3 shows winning probabilities (P), payoffs (V), and expected values (EV) for the three choice alternatives W, L, and S of each of Hommers' 15 stimuli. Stimuli were presented to Ss in form of index cards showing sets of "winning" and "not winning" balls in urns, and displaying the amounts to be won in coins. Subjects made their choice by indicating their favored gamble, which was played thereafter. About

half of the Ss in each age and school level had previous experience with choices on stimulus cards with two choice alternatives, so that there are three independent variables: school level (normal vs. educable retarded), age level, and prior gambling experience vs. no prior gambling experience.

Hommers' data, i.e., frequencies of choices of the alternatives W, L, and S of the 15 stimuli in the 14 groups, are displayed in Table 4. Hommers' analysis of these data consisted of chi square comparisons between chese figures, testing various hypotheses about differences in the development of risk vs. safety orientation and EV maximization between the age groups tested and between the normal and educable retarted children.

However, since it is assumed that these children follow some probabilistic choice model, it is feasible to apply a BTL choice model to these data, and do a likelihood ratio analysis. Three probabilistic choice models derived from Hommers' hypotheses seem to be naturally applicable in this situation: Ss are either (1) safety oriented, i.e., focussing on the probability of winning, and thus should choose the alternatives with probabilities proportional to their respective winning probabilities, or (2) they are value oriented, and choose with probabilities proportional to the payoffs, or (3) they are expected-value oriented, and choose with probabilities proportional to the expected values of the alternatives. All wins and expected values are positive. Choice probabilities for the alternatives W, L, and S of each stimulus are calculated under the assumption of each of these three models, and displayed in Table 5. In these computations, use has been made of the "auxillary sums" in the last three columns of Table 3; e.g., in stimulus 1, the sum of the EV

Table 3: Hommers' (1973) stimuli: three-alternative choices among bets

stimulus #	alternative (W) P	W	EV	alternative (L) P	W	EV	alternative (S) P	W	EV	auxiliary sums Σ P	Σ W	Σ EV
1	.1	15	1.5	.5	10	5.0	.9	5	4.5	1.5	30	11.0
2	.3	35	10.5	.5	15	7.5	.7	10	7.0	1.5	60	25.0
3	.1	25	2.5	.3	15	4.5	.5	10	2.5	.9	50	9.5
4	.1	15	1.5	.7	10	7.0	.9	5	4.5	1.7	30	13.0
5	.1	35	3.5	.3	25	7.5	.5	15	7.5	.9	75	18.5
6	.1	35	3.5	.3	10	3.0	.7	5	3.5	1.1	50	10.0
7	.3	15	4.5	.5	10	5.0	.7	5	3.5	1.5	30	13.0
8	.3	35	10.5	.5	20	10.0	.7	15	10.5	1.5	70	31.0
9	.5	35	17.5	.7	25	17.5	.9	15	13.5	2.1	75	48.5
10	.3	25	7.5	.5	15	7.5	.9	10	9.0	1.7	50	24.0
11	.3	35	10.5	.5	25	12.5	.7	15	10.5	1.5	75	33.5
12	.3	30	9.0	.7	20	14.0	.9	10	9.0	1.9	60	32.0
13	.3	15	4.5	.5	10	5.0	.9	5	4.5	1.7	30	14.0
14	.5	25	12.5	.7	15	10.5	.9	5	4.5	2.1	45	27.5
15	.1	15	1.5	.3	10	3.0	.9	5	4.5	1.3	30	9.0

Note: maximal EV underlined; by dashed line where 2 maxima

Table 4: Hommer's data: absolute choice frequencies in

groups without prior gambling experience

stimulus #	V 8 n = 15 W	L	S	V 10 n = 15 W	L	S	V 12 n = 10 W	L	S	S 8 n = 8 W	L	S	S 10 n = 16 W	L	S	S 12 n = 13 W	L	S	S 14 n = 18 W	L	S
1	6	1	8	3	4	8	1	4	5	3	0	5	7	0	9	5	4	4	2	5	11
2	8	0	7	4	4	7	3	2	5	4	1	3	10	1	5	8	0	5	6	1	11
3	5	2	8	2	0	13	0	4	6	4	0	4	5	1	10	6	0	7	5	3	10
4	3	4	8	1	10	4	0	7	3	4	1	3	4	5	7	0	5	8	3	8	7
5	4	2	9	1	0	14	1	3	6	3	0	5	5	0	11	2	2	9	4	1	13
6	4	2	9	2	4	9	3	3	4	3	0	5	5	3	8	2	2	9	2	4	12
7	5	3	7	1	9	5	1	5	4	3	2	3	5	5	6	2	2	9	4	9	5
8	5	2	8	3	2	10	2	3	5	4	2	2	6	3	7	2	2	9	5	3	10
9	5	3	7	8	3	4	6	3	1	4	0	4	8	1	7	3	1	9	8	2	8
10	5	2	8	1	6	8	2	4	4	3	1	4	5	4	7	1	3	9	3	3	12
11	3	3	9	1	6	8	1	7	2	3	2	3	5	2	9	3	1	9	4	4	10
12	4	4	7	1	12	2	2	7	1	4	1	3	5	5	6	3	2	8	4	5	9
13	3	3	9	1	9	5	1	6	3	3	2	3	8	3	5	4	3	6	3	8	7
14	7	1	7	8	3	4	8	1	1	5	1	2	11	1	4	7	2	4	11	3	4
15	4	2	9	2	5	8	0	5	5	4	1	3	5	2	9	1	4	8	3	3	12

groups with prior gambling experience

stimulus #	n = 15 W	L	S	n = 15 W	L	S	n = 9 W	L	S	n = 6 W	L	S	n = 15 W	L	S	n = 17 W	L	S	n = 16 W	L	S
1	6	6	3	1	3	11	1	6	2	0	3	3	2	7	6	1	8	8	3	5	8
2	9	1	5	6	1	8	5	1	3	2	0	4	2	3	10	9	6	2	4	3	9
3	5	4	6	0	7	18	2	4	3	1	0	5	2	5	8	2	5	10	1	3	12
4	3	7	5	0	10	5	1	6	2	0	2	4	3	6	6	0	13	4	0	9	7
5	6	[illegible]	6	1	2	12	1	3	5	2	0	4	5	6	4	1	4	12	1	3	12
6	6	2	7	2	1	12	4	0	5	2	0	4	4	2	9	1	6	10	2	4	10
7	7	4	4	3	7	5	3	5	1	2	1	3	5	1	7	4	11	2	3	7	6
8	7	2	6	3	4	8	6	2	1	2	0	4	8	1	6	7	6	4	5	3	8
9	11	1	3	8	4	3	6	0	3	2	3	1	8	2	5	11	3	3	7	4	5
10	5	1	9	7	3	5	4	3	2	3	2	1	4	1	10	5	5	7	2	6	8
11	5	3	7	7	3	5	3	5	1	2	1	3	7	1	7	3	7	7	2	7	7
12	7	4	4	3	7	5	4	5	0	3	2	1	8	3	4	2	11	4	1	10	5
13	5	3	7	3	8	4	3	4	2	2	3	1	6	4	5	3	10	4	4	6	6
14	10	0	5	8	4	3	6	3	0	3	3	0	9	2	4	10	5	2	8	4	4
15	5	3	7	3	6	6	3	1	5	2	0	4	4	4	7	2	4	11	2	4	10

of the three-choice alternatives is 11.0 (= 1.5 + 5.0 + 4.5), and thus, under assumption of EV orientation, the choice probabilities of alternatives W, L, and S are 1.5/11.0 = .136, 5.0/11.0 = .455, and 4.5/11.0 = .409, respectively.

For convenience, the choice probabilities have been converted into logarithms in the right half of Table 5. As in the previous examples, we again assume independence of observations, so that the likelihood of the whole set of data (observed choice frequencies) or of parts thereof is equal to the product of choice probabilities under assumption of the various models. In logarithms, this means multiplying the choice frequencies from Table 4 to the logarithms of choice probabilities from Table 5, and then summing up over alternatives and stimuli for each model. The antilog of this sum is the likelihood of the data set under the specified hypothesis or model. These likelihoods can be compared pairwise between models (but only for the same data set); however, the resulting likelihood ratios can be compared between data sets, i.e., between the different experimental groups.

For some of Hommers' (1973) data, this has been done in Tables 6-9. The sume in the bottom rows are the logarithms of the likelihoods (probabilities) of the respective data, assuming that the probabilities of individual choices are generated by the models named on top of the columns. Of course, they are all negative; the larger their absolute value, the smaller the probability of the data under the respective model.

In the order of their likelihoods, we get from the four groups analyzed the following likelihood ratios between pairs of models (see Table 10).

Table 5: Choice probabilities from probabilistic choice models

BTL choice probabilities assuming								
focussing P			focussing V			focussing EV		
W	L	S	W	L	S	W	L	S
.067	.333	.600	.500	.333	.167	.136	.455	.409
.200	.333	.467	.583	.250	.167	.420	.300	.280
.111	.333	.556	.500	.300	.200	.263	.474	.263
.059	.411	.530	.500	.333	.167	.115	.539	.346
.111	.333	.556	.467	.333	.200	.189	.405	.405
.091	.273	.636	.700	.200	.100	.350	.300	.350
.200	.333	.467	.500	.333	.167	.346	.384	.270
.200	.333	.467	.500	.286	.214	.339	.322	.339
.238	.333	.429	.467	.333	.200	.361	.361	.278
.177	.294	.529	.500	.300	.200	.312	.312	.376
.200	.333	.467	.467	.333	.200	.313	.374	.313
.158	.369	.473	.500	.333	.167	.281	.438	.281
.176	.294	.530	.500	.333	.167	.322	.355	.322
.238	.333	.429	.556	.333	.111	.455	.381	.164
.077	.231	.692	.500	.333	.169	.167	.333	.500

logarithms of choice probabilities								
focussing P			focussing V			focussing EV		
W	L	S	W	L	S	W	L	S
-1.1739	-.4776	-.2218	-.3010	-.4776	- .7773	-.[illegible]5	-.3420	-.3[illegible]
- .6990	-.4776	-.3307	-.23-3	-.6021	- .7773	-.3[illegible]	-.5229	-.5[illegible]
- .9547	-.4776	-.2549	-.3010	-.5229	- .6990	-.5800	-.3242	-.5[illegible]
-1.2291	-.3862	-.2757	-.3010	-.4776	- .7773	-.9393	-.2[illegible]	-.[illegible]
- .9547	-.4776	-.2549	-.3307	-.4776	- .6990	-.7235	-.3925	-.3925
-1.0410	-.5638	-.1965	-.1549	-.6990	-1.0000	-.4559	-.5229	-.4559
- .6990	-.4776	-.3307	-.3010	-.4776	- .7773	-.4609	-.4157	-.5[illegible]
- .6990	-.4776	-.3307	-.3010	-.5436	- .6696	-.4698	-.4921	-.4[illegible]
- .6234	-.4776	-.3675	-.3307	-.4776	- .6990	-.4425	-.4425	-.55[illegible]
- .7520	-.5317	-.2765	-.3010	-.5229	- .6990	-.5058	-.5058	-.4[illegible]
- .6990	-.4776	-.3307	-.3307	-.4776	- .6990	-.5045	-.4271	-.5[illegible]
- .8013	-.4330	-.3251	-.3010	-.4776	- .7773	-.5513	-.3595	-.5513
- .7545	-.5317	-.2757	-.3010	-.4776	- .7773	-.4921	-.448[illegible]	-.4921
- .6234	-.4776	-.3675	-.2549	-.4776	- .9547	-.3420	-.4101	-.[illegible]52
-1.1135	-.6364	-.1599	-.3010	-.4776	- .7773	-.7773	-.4776	-.3[illegible]

$\Sigma(S) = -4.2988$ $\quad -4.3452 = \Sigma(W)$ $\quad \Sigma(EV) = -5.7898$

Table 6: Data from group V 8 o

stimulus #	focussing V						focussing EV						focussing P					
	W		L		S		W		L		S		W		L		S	
	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P
1	6	-.3010	1	-.4776	8	- .7773	6	-.8665	1	-.3420	8	-.3883	6	-1.1739	1	-.4776	8	-.2218
2	8	-.2343	0	-.6021	7	- .7773	8	-.3768	0	-.5229	7	-.5528	8	- .6990	0	-.4776	7	[illegible]
3	5	-.3010	2	-.5229	8	- .6990	5	-.5800	2	-.3242	8	-.5800	5	- .9547	2	-.4776	8	-.2840
4	3	-.3010	4	-.4776	8	- .7773	3	-.9393	4	-.2784	8	-.4609	3	-1.2201	4	-.3802	8	-.2757
5	4	-.3307	2	-.4776	9	- .6990	4	-.7235	2	-.3925	9	-.3925	4	- .9547	2	-.4776	9	-.2840
6	4	-.1549	2	-.6990	9	-1.0000	4	-.4559	2	-.5229	9	-.4559	4	-1.0410	2	-.5638	9	-.1965
7	5	-.3010	3	-.4776	7	- .7773	5	-.4609	3	-.4157	7	-.5686	5	- .6990	3	-.4776	7	-.3307
8	5	-.3010	2	-.5436	8	- .6696	5	-.4698	2	-.4921	8	-.4698	5	- .6990	2	-.4776	8	-.3307
9	5	-.3307	3	-.4776	7	- .6990	5	-.4425	3	-.4425	7	-.5560	5	- .6234	3	-.4776	7	-.3675
10	5	-.3010	2	-.5229	8	- .6990	5	-.5058	2	-.5058	8	-.4248	5	- .7520	2	-.5317	8	-.2765
11	3	-.3307	3	-.4776	9	- .6990	3	-.5045	3	-.4271	9	-.5045	3	- .6990	3	-.4776	9	-.3307
12	4	-.3010	4	-.4776	7	- .7773	4	-.5513	4	-.3585	7	-.5513	4	- .8013	4	-.4330	7	-.3251
13	3	-.3010	3	-.4776	9	- .7773	3	-.4921	3	-.4486	9	-.4921	3	- .7545	3	-.5317	9	-.2757
14	7	-.2549	1	-.4776	7	- .9547	7	-.3420	1	-.4191	7	-.7852	7	- .6234	7	-.4776	7	-.3675
15	4	-.3010	2	-.4776	9	- .7773	4	-.7773	2	-.4776	9	-.3010	4	-1.1135	2	-.6374	9	-.1599
	-129.7313						-111.6444						-109.6710					

log LR (focussing EV/focussing V) = -111.6444 - (-129.7313) = 18.0869 ⟶ LR = $1.221 \cdot 10^{18}$

log LR (focussing P/focussing EV) = -109.6710 - (-111.6444) = 1.9734 ⟶ LR = 94.06

log LR (focussing P/focussing V) = -109.6710 - (-129.7313) = 20.0603 ⟶ LR = $1.149 \cdot 10^{20}$

Table 7: Data from group S & o

focussing V						focussing EV						focussing P					
W		L		S		W		L		S		W		L		S	
choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P
3	-.3010	0	-.4776	5	-.7773	3	-.5665	0	-.3420	5	-.3883	3	-1.1730	0	[illegible]	5	[illegible]
4	-.2343	1	-.6021	3	-.7773	4	-.3768	1	-.5229	3	-.5528	4	[illegible]	1	[illegible]	3	[illegible]
4	-.3010	0	-.5229	4	-.6990	4	-.5800	0	-.3242	4	-.5800	4	[illegible]	0	-.4776	4	[illegible]
4	-.3010	1	-.4776	3	-.7773	4	-.9393	1	-.2684	3	-.4609	4	-1.2291	1	-.3562	3	[illegible]
3	-.3307	0	-.4776	5	-.6990	3	-.7235	0	-.3925	5	-.3925	3	-.9547	0	-.4776	5	[illegible]
3	-.1549	0	-.6990	5	-1.0000	3	-.4559	0	-.5229	5	-.4559	3	-1.0410	0	-.5638	5	-.1065
3	-.3010	2	-.4776	3	-.7773	3	-.4609	2	-.4157	3	-.5686	3	-.6990	2	-.4776	3	[illegible]
4	-.3010	2	-.5436	2	-.6696	4	-.4698	2	-.4921	2	-.4698	4	-.6990	2	-.4776	2	[illegible]
4	-.3307	0	-.4776	4	-.6990	4	-.4425	0	-.4425	4	-.5560	4	-.6234	0	-.4776	4	-.3675
3	-.3010	1	-.5229	4	-.6990	3	-.5058	1	-.5058	4	-.4248	3	-.7520	1	-.5317	4	[illegible]
3	-.3307	2	-.4776	3	-.6990	3	-.5045	2	-.4271	3	-.5045	3	-.6090	2	-.4776	3	[illegible]
4	-.3010	1	-.4775	3	-.7773	4	-.5513	1	-.3585	3	-.5513	4	-.8013	1	-.4330	3	[illegible]
3	-.3010	2	-.4776	3	-.7773	3	-.4921	2	-.4486	3	-.4921	3	-.7545	2	-.5317	3	[illegible]
5	-.2549	1	-.4776	2	-.9547	5	-.3420	1	-.4191	2	-.7852	5	-.6234	1	-.4776	2	[illegible]
4	-.3010	1	-.4776	3	-.7773	4	-.7773	1	-.4776	3	-.3010	4	[illegible]	1	-.6364	3	[illegible]
-62.7101						-61.6352						-67.1343					

log LR (focussing EV/focussing V) = -61.6352 - (-62.7101) = 1.0749 ⟶ LR = 11.88

log LR (focussing EV/focussing P) = -61.6352 - (-67.1343) = 5.4991 ⟶ LR = $3.156 \cdot 10^5$

log LR (focussing V/focussing P) = -62.7101 - (-67.1343) = 4.4242 ⟶ LR = $2.656 \cdot 10^4$

Table 8: Data from group S 12 o

focussing V						focussing EV						focussing P					
W		L		S		W		L		S		W		L		S	
choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P
5	-.3010	4	-.4776	4	- .7773	5	-.6665	4	-.3420	4	-.3883	5	-1.1739	4	-.4776	4	-.2218
8	-.2343	0	-.6021	5	- .7773	8	-.3768	0	-.5229	5	-.5528	8	- .6990	0	-.4776	5	-.3307
6	-.3010	0	-.5228	7	- .6990	6	-.5800	0	-.3242	7	-.5800	6	- .9547	0	-.4776	7	-.2549
0	-.3010	5	-.4776	8	- .7773	0	-.9393	5	-.2684	8	-.4609	0	-1.2291	5	-.3862	8	-.2757
2	-.3307	2	-.4776	9	- .6990	2	-.7235	2	-.3925	9	-.3925	2	- .9547	2	-.4776	9	-.2549
2	-.1549	2	-.6990	9	-1.0000	2	-.4559	2	-.5229	9	-.4559	2	-1.0410	2	-.5638	9	-.1965
2	-.3010	2	-.4776	9	- .7773	2	-.4609	2	-.4157	9	-.5686	2	- .6990	2	-.4776	9	-.3307
2	-.3010	2	-.5436	9	- .6690	2	-.4698	2	-.4921	9	-.4698	2	- .6990	2	-.4776	9	-.3307
3	-.3307	1	-.4776	9	- .6990	3	-.4425	1	-.4425	9	-.5560	3	- .6234	1	-.4776	9	-.3675
1	-.3010	3	-.5228	9	- .6990	1	-.5058	3	-.5058	9	-.4248	1	- .7520	3	-.5317	9	-.2765
3	-.3307	1	-.4776	9	- .6990	3	-.5045	1	-.4271	9	-.5045	3	- .6990	1	-.4776	9	-.3307
3	-.3010	2	-.4776	8	- .7773	3	-.5513	2	-.3585	8	-.5513	3	- .8013	2	-.4330	8	-.3251
4	-.3010	3	-.4776	6	- .7773	4	-.4921	3	-.4486	6	-.4921	4	- .7545	3	-.5317	6	-.2757
7	-.2549	2	-.4776	4	- .9547	7	-.3420	2	-.4191	4	-.7852	7	- .6234	2	-.4776	4	-.3675
1	-.3010	4	-.4776	8	- .7773	1	-.7773	4	-.4776	8	-.3010	1	-1.1135	4	-.6364	8	-.1599
-116.5222						-94.0674						-88.2646					

log LR (focussing P/focussing EV) = -88.2646 - (- 94.0674) = 5.8028 ⟶ LR = $6.35 \cdot 10^5$

log LR (focussing P/focussing V) = -88.2646 - (-116.5222) = 28.2576 ⟶ LR = $1.785 \cdot 10^{28}$

log LR (focussing EV/focussing V) = -94.0674 - (-116.5222) = 22.4548 ⟶ LR = $2.85 \cdot 10^{22}$

Table 9: Data from group S 12 m

focussing V						focussing EV						focussing P					
W		L		S		W		L		S		W		L		S	
choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P	choices	log P
1	-.3010	8	-.4776	8	- .7773	1	-.8665	8	-.3420	8	-.3883	1	-1.1739	8	-.4776	8	-.2218
9	-.2343	6	-.6021	2	- .7773	9	-.3768	6	-.5229	2	-.5528	9	- .6990	6	-.4776	2	-.3307
2	-.3010	5	-.5229	10	-. 6990	2	-.5800	5	-.3242	10	-.5800	2	- .9547	5	-.4776	10	-.2549
0	-.3010	13	-.4776	4	- .7773	0	-.9393	13	-.2684	4	-.4609	0	-1.2291	13	-.3862	4	-.2757
1	-.3307	4	-.4776	12	- .6990	1	-.7235	4	-.3925	12	-.3925	1	- .9547	4	-.4776	12	-.2549
1	-.1549	6	-.6990	10	-1.0000	1	-.4559	6	-.5229	10	-.4559	1	-1.0410	6	-.5633	10	-.1965
4	-.3010	11	-.4776	2	- .7773	4	-.4609	11	-.4157	2	-.5666	4	- .6990	11	-.4776	2	-.3307
7	-.3010	6	-.5436	4	- .1696	7	-.4198	6	-.4921	4	-.4698	7	- .6990	6	-.4776	-	-.3307
11	-.3307	3	-.4776	3	- .6990	11	-.4425	3	-.4425	3	-.5560	11	- .6234	3	-.4776	3	-.3675
5	-.3010	5	-.5229	7	- .6990	5	-.5058	5	-.5058	7	-.4248	5	- .7520	5	-.5317	7	-.2765
3	-.3307	7	-.4776	7	- .6990	3	-.5045	7	-.4271	7	-.5045	3	- .6990	7	-.4776	7	-.3307
2	-.3010	1	-.4776	4	- .7773	2	-.5513	11	-.3585	4	-.5513	2	- .8013	11	-.4330	4	-.3251
3	-.3010	10	-.4776	4	- .7773	3	-.4921	10	-.4486	4	-.4921	3	- .7545	10	-.5317	4	-.2757
10	-.2549	5	-.4776	2	- .9547	10	-.3420	5	-.4191	2	-.7852	10	- .6234	5	-.4776	2	-.3675
2	-.3010	4	-.4776	11	- .7773	2	-.7773	4	-.4776	11	-.3010	2	-1.1135	4	-.6344	11	-.1599
-139.2482						-112.0598						-117.4016					

$$\log LR\ (\text{focussing EV/focussing P}) = -112.0598 - (-117.4016) = 5.3418 \rightarrow LR = 2.197 \cdot 10^{5}$$

$$\log LR\ (\text{focussing EV/focussing V}) = -112.0598 - (-139.2482) = 27.1884 \rightarrow LR = 1.543 \cdot 10^{27}$$

$$\log LR\ (\text{focussing P/focussing V}) = -117.4016 - (-139.2482) = 21.8466 \rightarrow LR = 7.024 \cdot 10^{21}$$

Table 10: Examples of likelihood ratios from Hommer's data

Group 8 V o (8-year-old normal students without gambling experience):

likelihood ratio between:	more favored model— focussing P	more favored model— focussing EV	
less favored model—focussing EV	94.06		rank order of models:
focussing V	$1.149 * 10^{20}$	$1.221 * 10^{18}$	P - EV - V

Group 8 S o (8-year-old educable retarded children without gambling experience):

likelihood ratio between:	more favored model— focussing EV	more favored model— focussing V	
less favored model—focussing V	11.88		rank order of models:
focussing P	$3.156 * 10^{5}$	$2.656 * 10^{4}$	EV - V - P

Group 12 S o (12-year-old educable retarded children without gambling experience):

likelihood ratio between:	more favored model— focussing P	more favored model— focussing EV	
less favored model—focussing EV	$6.35 * 10^{5}$		rank order of models:
focussing V	$1.785 * 10^{28}$	$2.85 * 10^{22}$	P - EV - V

Group 12 S m (12-year-old educable retarded children with gambling experience):

likelihood ratio between:	more favored model— focussing EV	more favored model— focussing P	
less favored model—focussing P	$2.197 * 10^{5}$		rank order of models:
focussing V	$1.543 * 10^{27}$	$7.024 * 10^{21}$	EV - P - V

Similar analysis could be performed for other 10 of Hommers' 14 groups too. We have displayed in the rightmost column of Table 10 the rank order of models as indicated by the likelihood ratios calculated from the data; although the likelihood ratios themselves differ considerably, it is interesting to note that 12 year old retarded children show the same rank order of models as the 8 year old normal children, thus supporting Hommers' hypothesis of retardation as a shift in development. Also, comparison of the results from 12 year old educable retarded children without gambling experience with those from their classmates with prior gambling experience unveils a considerable influence of this experience on choices among gambles.

Besides these analyses for individual groups, larger groups can be taken into consideration, e.g., likelihood ratios between models can be calculated over all Ss with prior gambling experience, or over all retarded children to be compared to those calculated over all normal children, etc. Since we used these data only for illustrative purposes, we need not go into further detail. Also, we will turn to the problem of interpretation of such analyses later in this paper

Seghers, Fryback & Goodman's Data

The next set of data we are going to use are those of Seghers, Fryback & Goodman (1973). They presented their Ss sets of 7 gambles, like those reproduced in Table 11:

Table 11: List #1 as an example

bet #	win on 4	lose on 32	EV	Var
1	1.55	1.10	- .806	.683
2	3.45	1.15	- .639	2.088
3	5.30	1.20	- .478	4.469
4	7.15	1.25	- .317	6.963
5	8.95	1.30	- .162	10.423
6	10.80	1.35	0	14.567
7	12.65	1.40	+ .162	19.479

Wins and losses were determined by means of a roulette wheel which was respun if 0 or 00 occurred, such that "win on 4" (numbers) means a winning probability of 4/36 = 1/9, etc.

Seghers, Fryback & Goodman's lists varied in

(1) expected value (EV),

(2) range of outcomes (A-B),

(3) step size of expectation increase (ΔEV),

(4) position of the maximal EV bet (OBP).

Dependent variables were:

(a) choice of most perferred gamble.

(b) rank orderings of the sets of 7 gambles.

Although the experimental design looks as though a factorial design AVOVA had been planned, the data don't permit such an analysis. A frequency analysis as suggested by Sutcliffe (1957) would be more appropriate, however, low expected cell frequencies in the overall contingency table prohibits such an analysis.

A Bayesian data analysis is suggested as an alternative.

However, since Seghers, Fryback & Goodman assume a deterministic decision making model, this analysis runs into the problems mentioned before. The simple probabilistic choice model used to analyze Hommers' data is no longer appropriate here since there are negative expectations which are not compatible with a BTL choice model based on these expectations as scale values.

Deterministic decision making models predict choice of the optimal gamble with probability 1, and of all other alternatives with probability 0

$$P(\text{choice of gamble } g_j) = \begin{cases} 1 \text{ if } g_j \text{ is optimal} \\ 0 \text{ else} \end{cases}$$

where "optimal" is defined in the context of the respective decision making model to be tested, e.g., it would be the maximum EV bet under the expectation maximization model, or the ideal risk bet under assumption of Coombs Portifolio Theory. Unfortunately, likelihoods of 0 or 1 cannot be handled by the Bayesian data analysis model. Thus, we have to modify these models somehow to get away from the 0-1 likelihoods. There are several ways to do so of which we will try to

(1) keep the deterministic model in principle, but dilute the too peaked o-1 likelihood function by allowing for some error variance,

(2) modify the deterministic hypothesis somewhat arbitrarily to smooth its peak, following an example given by Pitz (1968), who encountered a similar problem,

(3) abandon the deterministic model completely in favor of some probabilistic choice model (as they have been used for riskless choices for a long time),

(4) replace the deterministic model by some hybrid of deterministic and probabilistic components.

We will explore all these possibilities in turn.

(1): Introducing error variance: Our suggestion is to dilute the too peaked likelihood functions somewhat by allowing for error variance: The diluted H_1 no longer assumes Ss always pick the maximal EV gamble, but rather assumes that Ss err sometimes in the sense that they don't choose a certain gamble although they mean to choose it.

Fortunately, the data by Seghers, Fryback & Goodman provide a way to estimate these error rates: they had their Ss do the task twice. Our suggestion is to use the observed discrepancies between first and second choice (under otherwise equal conditions) as estimates of error rates. To do so, the Ss first and second choices of gambles are tallied in 7x7 confusion matrices, separately for each given position of optimal EV bet (OBP). A completely consistent S should make the same choice on both occasions: i.e., all entries should be in the main diagonal, and all other cells should be empty. Every deviation from this diagonal matrix is considered an "error," an inconsistency, a deviation of the S from his pure strategy assumed under the hypothesis of

expectation maximization, H_1. Assuming that Ss err at both choices, i.e., both 1st and 2nd choices have a chance to deviate from the Ss' true choice predicted by his strategy, we take the average of row and column distribution for each stimulus as its error distribution.

This procedure assumes that, on the 2 days, S at least once chooses his "ideal bet" without making an error. It does not take into account those cases where S "wants to" select a certain bet but "misses" on both days. This may lead to an underestimation of error rates. A better way would be to get confusion probability estimates from more often repeated choices, in a complete pair comparison matrix, or from a different task, like the procedure used in DeSoto & Bosley (1962) (quoted in Coombs, Dawes & Tversky, 1970, p. 68 ff.). This cannot be done with these data, but it could be in future experiments— if you want to make the assumption that confusion of memory traces is representative of confusion in choices.

Now, with this knowledge about S's error probabilities, we can modify the 0-1 distribution under the former pure expectation maximization hypothesis: We diminish the peak of the distribution (formerly $P(D|H_1) = 1$ at maximal EV bet) by replacing the 1 by the repetition rate (1st choice = 2nd choice) in 1st choice/2nd choice confusion matrix, and by replacing the zeroes by the relative frequencies with which Ss have chosen the respective gambles "erroneously."

Thus, the EV maximization hypothesis H_1 implies data probabilities of

$P(D_o|H_1)$ = the repetition probability of the maximal EV bet for the maximal EV bet (D_o) chosen

and

$P(D_i | H_1)$ = the probability of choosing D_i given S has chosen D_o on
$i \neq o$ the same trial in the 1st or 2nd repetition.

($\sum_i P(D_i | H_1)$ should be 1 if everything is correct.) Analogous computations can be done for other alternative hypotheses, like variance perference, also.

Tables 12 and 13 give examples of such confusion matrices between 1st and 2nd choice: Table 12 are absolute frequencies; Table 5 is the same matrix with a matrix of ones added to it. (Actually, the entries in Table 12 are averaged over 2 presentations.)

The rationale for adding these ones to the cells is again a Bayesian one: we are revising here, in principle, Dirichlet distributions (see, e.g., Novick & Grizzle, 1965). We start with a uniform (flat) prior distribution D(1, 1, 1, 1, 1, 1, 1) with all parameters equal to 1, and then add to them the numbers of observations to obtain the parameters of the posterior distribution after Bayesian revision. However, summing cell entries from row and column would assume independence of observations from the two sessions which probably is not given since we assume that S's choices were influenced by the same preference structure on both days. Thus, to avoid an overly peaked Dirichlet distribution, we average over column and row entry rather than adding them up. Actually, this does not make a difference as long as we calculate only means and not variances.

Table 12: Choice on day 2/choice on day 1

averaged confusion matrix $\frac{G + R_1 0}{2}$ V $\underline{S}$s

Overall opt. bets	1	2	3	4	5	6	7	
1	116.5	12.5	7.5	2.5	1.5	3	5	148.5
2	13.5	13.5	9.5	6	1.5	0	1.5	45.5
3	9.5	7	25.5	6	2.5	0.5	3	54
4	5	1	7.5	6	2	2	2.5	26
5	3.5	1	1	6.5	16.5	1	3	32.5
6	1	1	2	2.5	3	0.5	4	14
7	8	2.5	5	4	5.5	1.5	37	63.5
	157	38.5	58	33.5	32.5	8.5	56	384

Table 13: Matrix with 1 added to every cell

+ 1 in all cells	1	2	3	4	5	6	7	
1	117.5	13.5	8.5	3.5	2.5	4	6	155.5
2	14.5	14.5	10.5	7	2.5	1	2.5	52.5
3	10.5	8	26.5	7	3.5	1.5	4	61
4	6	2	8.5	7	3	3	3.5	33
5	4.5	2	2	7.5	17.5	2	4	39.5
6	2	2	3	3.5	4	1.5	5	21
7	9	3.5	6	5	6.5	2.5	38	70.5
	164	45.5	65	40.5	39.5	15.5	63	433

As an illustration, assuming that gamble #1 is the optimal bet in the Ss' view (H_2), and having observed the number of choices displayed in Table 13, we get:

Table 14

from row 1 :	117.5	13.5	8.5	3.5	2.5	4	6
from column 1 :	117.5	14.5	10.5	6	4.5	2	9
sum of both :	235	28	19	9.5	7	6	15
average :	117.5	14	9.5	4.25	3.5	3	7.5
and thus the choice probabilities :	.734	.088	.060	.027	.022	.019	.047
for gamble # :	1	2	3	4	5	6	7

when gamble #1 is the "true choice" assumed by the model.

Some results of such tallies are reproduced in Table 15, assuming various choice strategies on the side of the Ss. Column 2 displays choice probabilities under an a priori random-choice null hypothesis (all gambles chosen with equal probability 1/7 = .143).

Table 15

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
gamble #	H_0: random choice	H_1: maximize EV: maximal EV is in gamble				H_2: always pick #1	H_3: always pick #1-3	H_4: always pick #5-7	H_5: always pick #3-5
		#1	#3	#5	#7				
1	.143	.802	.110	.080	.092	.734		.112	.127
2	.143	.060	.140	.051	.040	.088	.818	.054	.116
3	.143	.038	.566	.058	.046	.060		.080	
4	.143	.031	.065	.124	.050	.030	.063	.117	.594
5	.143	.019	.024	.482	.040	.022	.031		
6	.143	.018	.035	.082	.062	.019	.025	.652	.062
7	.143	.032	.060	.117	.670	.047	.057		.105

Columns 3 through 6 are the diluted choice probabilities assuming expectation maximization with some errors, calculated in the manner described above from confusion matrices between choices in first and second sessions of Ss but tallied separately for lists where gambles 1, 3, 5, and 7 were optimal, respectively.

Column 7 is calculated from the tallies illustrated in Tables 12, 13, and 14, assuming that Ss have the strategy of always picking gamble #1, no matter what the parameters of the gambles in the list are.

Columns 8 through 10 are choice probabilities calculated under similar hypotheses, assuming that Ss have preferences for certain regions of the lists of gambles presented to them, i.e., that they always pick gambles #1-3, or #5-7, or #3-5, respectively.

With the choice probabilities from Table 15 taken as $P(D|H_i)$, all these models can be tested against each other by calculating the respective likelihood ratios. To make the analysis more convenient, all hypotheses could be tested first against the random-choice null hypothesis (H_o). The resulting likelihood ratios against H_o could then be divided by each other to yield likelihood ratios agains each other since

$$\frac{P(D|H_i)}{P(D|H_o)} \Big/ \frac{P(D|H_j)}{P(D|H_o)} = \frac{P(D|H_i)}{P(D|H_j)}$$

However, this is only feasible as far as H_i and H_j are mutually exclusive. H_1, H_2 and H_3 in Table 15 are not since they all assume a strategy to choose gamble #1.

The choice probabilities assumed under hypotheses H_1 through H_5 from Table 15 yield the likelihood ratios reproduced in Table 16 if tested against the uniform distribution H_o.

To use Table 16, we multiply the entries by the prior odds every time the respective datum comes up; e.g., to test hypothesis H_1 against H_o, we would multiply prior odds (i.e., odds so far obtained) by 5.14 if S chooses gamble #1, and gamble #1 is optimal (maximal EV) in the respective list.

Table 16: Likelihood ratios calculated from Table 15

(1) Gamble #	(2) $LR_{1/0}$ 1 opt	(3) $LR_{1/0}$ 3 opt	(4) $LR_{1/0}$ 5 opt	(5) $LR_{1/0}$ 7 opt	(6) $LR_{2/0}$	(7) $LR_{3/0}$	(8) $LR_{4/0}$	(9) $LR_{5/0}$
1	5.61	.77	.56	.64	5.14		.78	.89
2	.42	.98	.36	.28	.62	1.91	.38	.81
3	.27	3.96	.41	.32	.42		.56	
4	.22	.46	.87	.35	.21	.44	.82	1.39
5	.13	.17	3.37	.28	.15	.22		
6	.13	.25	.57	.43	.13	.18	1.19	.43
7	.22	.42	.82	4.69	.33	.40		.74

Again, it will be more convenient to do this in terms of logarithms, thus we have, in Table 17, the log $LR_{1/0}$ in column 3, and the number of choices for the respective gamble in column 2.

Table 17

(1) gamble #	(2) number of choices	(3) log $LR_{1/0}$	(4) log $LR_{2/0}$	(5) log $LR_{4/0}$
1	3	- .1938	+ .7110	- .1079
2	0	- .5528	- .2076	- .4202
3	2	- .4949	- .3768	- .2518
4	2	- .4559	- .6778	- .0862
5	1	- .5528	- .8239	+ .0755
6	1	- .3665	- .8861	+ .0755
7	15	+ .6712	- .4815	+ .0755
log LR		+6.6657	- 8.9087	+ .2838
LR		$4.631*10^6$	$1/(8.104*10^8)$	1.922

The data in column 2 are the choices made by 12 Ss in 2 sessions among the gambles of list #1, reproduced in Table 11, where gamble #7 had maximal EV, such that the logarithms in column 3 of Table 17 are those of the likelihood ratios in column 5 or Table 16. The sum of the products of entries in columns 2 and 3 of Table 17, the overall log likelihood ratio, is 6.6657, indicating a likelihood ratio of $4.631*10^6$ in favor of expectation maximization (H_1) over random choice (H_o).

Columns 4 and 5 show the respective log LR for hypothesis H_2 (always pick gamble #1) over the random choice hypothesis H_o, and for hypothesis H_4 (always pick gamble # 5, 6, or 7) against the random choice hypothesis H_o. Resulting likelihood ratios $LR_{0/2} = 8.104*10^8$ in favor of H_o (random choice) over H_2

(always pick gamble #1) with these data, and $LR_{4/0}$ = 1.922 in favor of H_4 (always pick # 5, 6, or 7) over H_o (random choice).

So far, we have analyzed only the choices among gambles of one list— of course, it is feasible and advisible to do it over the whole set of data from all lists, simply by summing up the respective log $LR_{1/0}$ over all data for the various hypotheses H_i. Seghers, Fryback & Goodman have done this for each of their Ss, individually, and we are reproducing their results for one of their Ss as an example in Table 18. Besides calculating likelihood ratios $LR_{1/0}$ for the aforementioned hypotheses H_i against the random choice hypothesis H_o over all (lists) (column 2), they also did it for specified subsets of lists, e.g., lists with high EV (column 2), lists with low EV (column 4), lists with high EV differences between gambles in the lists (column 5), lists with low EV differences (column 6), lists of gambles with large variances (range of bet, i.e., |win-loss|) (column 7), and lists of gambles with small variances (column 8). Thus, it is possible to compare data likelihood, for the various hypotheses H_i under different stimulus conditions.

This breaking down likelihood ratio analyses into analyses over mutually exclusive subsets of the whole data set corresponds roughly to what is done to the sum of squares in analysis of variance (ANOVA), or to the chi square in analyses of multi-dimensional contingency tables (e.g., see Sutcliffe, 1957): It shows how much the respective subsets of data (i.e., data under specific conditions) contribute to the overall likelihood ratio. To make fair comparisons of this kind, we have to take care that these subsets are of equal size.

Table 18: Likelihood ratios for S #1 of Seghers, Fryback & Goodman

(1) competing hypotheses	(2) all lists	(3) only high EV lists	(4) only low EV lists	(5) only high EV difference lists	(6) only low EV difference lists	(7) only large range lists	(8) only small range lists
	LR calculated over:						
$LR_{1/0}$: expectation (EV) maximization vs. random choice	2867.3	33.5	85.6	262.2	10.9	446.6	6.4
$LR_{2/0}$: always pick #1 vs. random choice	3715.4	2752.6	1.4	109.7	33.9	361541.3	1/97.3
$LR_{3/0}$: always pick #1,2,3 vs. random choice	622.2	136.3	4.6	66.6	9.3	302.8	2.1
$LR_{4/0}$: always pick #5,6,7 vs. random choice	1/19743.6	262.2	75.3	170.8	115.6	62.2	317.2
$LR_{5/0}$: always pick #3,4,5 vs. random choice	1/2.7	1/7.9	2.9	1/2.9	1.1	1/3.2	1.2

Note: reciprocal values (1/x) indicate that the data were, in these cases, more likely under H_o than under H_i

The product of the likelihood ratios $LR_{i/j}$ competing hypotheses H_i, H_j from exhaustive and mutually exclusive subsets of data equals their likelihood ratio over the whole data set. E.g., in each row of Table 18, the products of entries in columns 3 and 4, 5 and 6, or 7 and 8 equal each other, and equal the entry of column 2, except for rounding errors. (This provides, by the way, an easy means of checking computations.)

The results of such likelihood ratio analyses over the subsets of data can be used to find out under which conditions which hypotheses are how much more likely than others, and thus may lead to more specific theories about the underlying pattern of behavior.

The comparison of likelihood ratio analysis to more conventional methods like ANOVA is not always straightforward; the easiest comparable traditional technique would be a frequency analysis because it deals with the frequencies of occurrence of events which enter directly the likelihood ratio analysis (as exponents.)

Seghers, Fryback & Goodman did analyses of variance over the same data we used for demonstration in Table 18, both terms of absolute deviation of bet number as dependent variable, and in terms of absolute deviation of bet number as dependent variable, and in terms of absolute deviation of bet number chosen from maximal EV bet number in the respective list. Results (for the same S, and same session as in Table 18) are shown in Table 19.

Seghers, Fryback & Goodman's lists were constructed in such a way that, given the maximal EV bet in the list (in positions #1, #3, #5, or #7 of the list = optimal bet position OBP), the adjacent gambles decreased in EV to both

Table 19: Analyses of variance for choices of S #1 of Seghers, Fryback & Goodman

source of variation	df	ANOVA of absolute deviation of bet chosen from maximal EV bet				ANOVA of absolute number of bet chosen			
		Σx^2	mean square	F-ratio if > 1	% variance accounted for	Σx^2	mean square	F-ratio if > 1	% variance accounted for
maximal EV (EV)	1	0	0			2.000	2.000		
EV difference (DEV)	1	2.000	2.000			0	0		
range (R)	1	0.125	0.125			3.125	3.125		
optimal bet portion (OBP)	3	45.625	15.208	4.31	26%	20.125	6.708	1.53	6%
interactions:									
EV x DEV	1	1.125	1.125			6.125	6.125	1.40	2%
EV x R	1	0	0			2	2		
EV x OBP	3	4.750	1.583			4.250	1.417		
DEV x R	1	4.500	4.500	1.27	1%	0.500	0.500		
DEV x OBP	3	5.750	1.917			8.250	2.750		
R x OBP	3	10.625	3.542			10.125	3.375		
EV x DEV x R	1	3.125	3.125			0.125	0.125		
EV x DEV x OBP	3	32.625	10.875	3.07	16%	28.125	9.375	2.14	13%
EV x R x OBP	3	6.750	2.250			5.250	1.750		
DEV x R x OBP	3	3.250	1.083			6.750	2.250		
residual (error)	3	10.625	3.542			13.125	4.375		
total	31	130.875				109.875			

sides by a step size DEV = difference in expected value. Thus, the dependent variable "absolute deviation of number bet chosen from number of maximal EV bet" can be considered a measure of S's deviation from expectation maximation behavior.

Whereas such independent variables like "high level of maximal EV in list" versus "low level of maximal EV in list" (first line in Table 19), large step size of EV differences in list versus small step size (line 2 in Table 19), and range of outcomes of gambles (line 3 in Table 19) show no significant difference in the dependent variables, there are some differences between the contributions of the respective subsets of data to the likelihood ratio between expectation maximization and random choice hypotheses in Table 18 (line 1). However, we have no means to compare these two kinds of analyses quantitatively.

Testing the various hypotheses H_i about choice behavior against the random choice hypothesis H_o is the approach to their evaluation that comes closest to traditional hypothesis testing. Testing them against the most descriptive choice probabilities is another possibility these likelihood analyses offer for which no counterpart exists in traditional statistics.

Comparisons of data likelihoods under the various hypotheses aforementioned to these (by definition) maximal likelihoods can show how far out hypotheses H_i deviate from actual behavior. These most descriptive choice probabilities specify upper bounds for data likelihoods, under the choice hypotheses, as illustrated in Figure 1.

L(D) ——┼————————┼————┼————┼————————————┼——

H_o normative models H optimal description

random choice, uniform p maximum likelihood p

Figure 1

The most descriptive (maximum likelihood) vector of choice probabilities for the seven gambles can be obtained for each subject from his choices by the following method: the data—choices of one out of seven gambles in each list—are generated by a multinomial distribution, with choice probabilities p_j following a Dirichlet distribution. Thus we can assume a flat Dirichlet distribution D(1, 1, 1, 1, 1, 1, 1) as prior, a multinomial data generating process yielding x_j choices of gamble g_i, and thus leading (via a Bayesian probability distribution revision) to a Dirichlet posterior distribution, $D(x_1 + 1, x_2 + 1, x_3 + 1, x_4 + 1, x_5 + 1, x_6 + 1, x_7 + 1)$. This Dirichlet posterior distribution gives us the probability $P(\bar{p}|\bar{x})$ of vector of choice probabilities $(p_1, p_2, p_3, p_4, p_5, p_6, p_7) = \bar{p}$ of gambles g_1 through g_7, given the vector of observed choice frequencies $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = \bar{x}$, and what we need is that vector $\bar{p}_o$ for which $P(\bar{p}|\bar{x})$ is maximal over the space of all possible $\bar{p}$. (Note that this space is restricted by $\sum_j p_j = 1$ for each $\bar{p}$.)

We take S #1 of Seghers, Fryback & Goodman, again, as an example. His (or, rather, her) choices are reproduced in columns 2, 5, 8, and 11 for the respective OBP conditions, and summed up in column 14 of Table 20. Columns 3, 6, 9, and 12 contain the choice probabilities under the diluted expectation maximization hypothesis H_1 from Table 15, in columns 4, 7, 10, and 13 we find the corresponding logarithms. The log likelihood for expectation maximization

Table 20

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
	choices of S #1 when optimal gamble was (expectation maximation hypothesis (H_1)												maximal descriptive strategy (H [illegible])				random choice (H [illegible])	
gamble #	gamble #1			gamble #3			gamble #5			gamble #7								
	choices	choice prob. (H_1)	log P	choices	choice prob. (H_1)	log P	choices	choice prob. (H_1)	log P	choices	choice prob. (H_1)	log P	total # of choices	# of choices +1	S's choice prob.	log P	choice prob.	log P
1	13	.802	- .0958	7	.110	- .9586	3	.080	-1.0969	5	.092	-1.0362	28	29	.400	- .3979	.143	-.8447
2	1	.060	-1.2218	4	.140	- .8539	3	.051	-1.2924	4	.040	-1.3979	12	13	.183	- .7375	.143	-.8447
3	1	.038	-1.4202	2	.566	- .2472	3	.058	-1.2366	2	.046	-1.3372	8	9	.127	- .8962	.143	-.8447
4	1	.031	-1.5086	3	.065	-1.1871	3	.124	- .9066	0	.050	-1.3010	7	8	.133	- .8760	.143	-.8447
5	0	.019	-1.7212	0	.024	-1.6198	4	.482	- .3170	0	.040	-1.3979	4	5	.070	-1.1549	.143	-.8447
6	0	.018	-1.7447	0	.035	-1.4559	0	.082	-1.0862	1	.062	-1.2076	1	2	.028	-1.5528	.143	-.8447
7	0	.032	-1.4949	0	.060	-1.2218	0	.117	- .9318	4	.670	- .1739	4	5	.070	-1.1549	.143	-.8447
log L													64	71				
						-49.7932									-44.3123			-54.0608

calculated from these figures is -49.7952. The $\underline{S}$'s most descriptive strategy, computed as outlined in the preceeding paragraph, is given in column 16, with the corresponding logarithms in column 17. The log likelihood from these figures (which is the maximal attainable) is -44.3123, and the log likelihood of this $\underline{S}$'s choices under the random choice hypothesis H_0 is $64 * \log 1/7 = -54.0608$. The expectation maximization hypothesis (H_1) comes much closer to the subjects most descriptive strategy (H_7) than to the random choice strategy (H_0). The respective likelihood ratios are

$$LR_{7/1} = 3.026 * 10^5 \qquad LR_{1/0} = 1.852 * 10^4$$

and

$$LR_{7/0} = 5.604 * 10^9$$

We have so far used the assumption that $\underline{S}$s occasionally deviate from their ideal choice and make "errors" in their decisions which we could use to get rid of the choice probabilities of 0 and 1 assumed by the deterministic normative models of decision making.

Expectation Preference Model

In discussing Hommers' paper, we have seen that the assumption of probabilistic preference models rather than deterministic choice models is another feasible way to avoid choice probabilities of 0 and 1.

For gambles of the form $g_j = (w_j, p_j, l_j)$ where Ss wins the payoff w_j with probability p_j and loses l_j with probability $(1-p_j)$, this model assumes that Ss choose a gamble g_j with probability $P(g_j)$ proportional to the relative utility $U(g_j)$ of the gamble g_j,

$$P(g_j) = U(g_j)/\sum_j U(g_j),$$

where

$$U(g_j) = EV(g_j) = p_j w_j + (1-p_j)l_j$$

under the expectation preference model. For each choice of g_j an S makes, $P(g_j)$ is the likelihood of this observation to occur under assumption of this model.

This expectation preference model works fairly well for sets of gambles where all EVs are positive, as we have seen in the analysis of Hommers' data. However, it will run into difficulties if the EV of one or more gambles in the list (set of choice alternatives) is negative or zero.

A Thurstonean (rather than Lucean) choice model might help in this case. Here, choice probabilities are only dependent on differences between utilities

of choice alternatives, and not on their absolute values. Under the assumptions of this model, the probability of choosing one element (i.e., a gamble) in a pair of alternatives is equal to the integral of the normal distribution from $-\infty$ to the difference in utilities (expected values) of the respective pair, where the mean of this normal distribution is 0, and its variance is the variance of the utility difference which is the sum of the variances of the discriminal dispersions of the two elements (gambles) in the pair, if we assume independence (uncorrelatedness) of these two discriminal processes. Application of this model requires estimation of these variances which can be obtained from repeated choices.

Regret Avoidance Models

A way to apply a Lucean choice model to choices among bets including gambles with $EV < 0$ might be to consider regrets rather than payoffs. Regrets are obtained from payoffs by reducing them by the maximal amount obtainable with each given state of world. Regrets calculated by this method are all negative; they are measures of undesirability rather than desirability. Thus, it does not make sense to assume choice probabilities proportional to regrets. What we need is some antitone transformation on the regrets which leads to high choice probabilities for low regrets, and low choice probabilities for large regrets. We propose three simple models for this purpose:

(a) the sum-difference regret model assumes that choice probabilities are proportional to the deviation of the respective expected regrets from the sum of all regrets,

$$P(i) = \frac{\sum_i r_i - r_i}{(N - 1) \sum_i r_i}$$

where r_i is the expected regret associated with the i^{th} alternative, smallest regret being 0, N=number of alternatives. Model (a) gives choice probabilities with a rather small variance, i.e., the choice probabilities are not very sensitive to differences in regrets.

(b) the <u>reciprocal regret model</u> assumes that choice probabilities are proportional to the reciprocals of the respective expected regrets,

$$P(i) = \frac{1}{r_i \sum_i \frac{1}{r_i}}$$

This leaves P(i) for $r_i = 0$ undefined. Model (b) leads to stronger deviations of choice probabilities from a uniform distribution over alternatives to differences in regrets, i.e., model (b) is more sensitive, but cannot always be used because if leaves the choice probability for an expected regret = 0 undefined.

(c) the <u>max-difference model</u> assumes that choice probabilities for alternatives i are proportional to the differences between the respective expected regrets and the maximal expected regret,

$$P(i) = \frac{\max_i [r_i] - r_i}{N \max_i [r_i] - \sum_{i=1}^{N} r_i}$$

This model is more sensitive to differences in expected regrets than model (a) and leaves no choice probabilities undefined as does model (b), but leads to a 0 choice probability for the maximal expected regret alternative. This is an undesirable consequence for a BTL choice model but may be quite

realistic. In the data analysis, it will hurt only if any S picks the maximum expected regret gamble.

For the example of list #1 from Seghers, Fryback & Goodman (see Table 11), Table 21 shows the respective choice probabilities with these probabilistic regret avoidance models in columns 8, 11, and 14, with the corresponding logarithms in columns 9, 12, and 15. Column 17 displays the choice probabilities under error-diluted deterministic expectation maximization hypothesis H_1 as given in Table 15, and column 18 of Table 21 contains their logarithms. In column 19, we have the actual numbers of choices made by S in this list of gambles, for which we calculated the likelihoods under the hypothesis H_o (random choice), H_1 (diluted expectation maximation), H_8 (reciprocal regret), H_9 (sum-difference regret), and H_{10} (max-difference regret). Table 22 displays the pairwise likelihood ratios between these hypotheses.

As we can see, the data are 1067 times more likely under the diluted deterministic expectation maximization hypothesis H_1 than under the most favored probabilistic regret-avoidance hypothesis H_8. The data likelihood under the least favored probabilistic regret-avoidance hypothesis H_9 is almost as large as under random choice assumption H_0, $LR_{9/0} = 1.111$.

This indicates that for likelihood ratio analyses of choices among bets made by adult subjects, error-diluted deterministic expectation maximization models seem much more likely than probabilistic preference models. However, in the case of Hommers' data where no source to estimate the error rate was available, probabilistic preference models proved quite useful. It should be mentioned that neither of these studies was originally designed for a likelihood ratio analysis—if this had been the case, adequate measures would

Table 21

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
gamble #	payoffs		regrets		expected regret (ER)	reciprocal regret model (H_8)			sum-diff. regret model (H_9)			max.-diff. regret model (H_{10})			expectation max. model (H_1)		number of choices made by S
	win	loss				1/ER	choice prob. (H_8)	log P	$\sum$ER-ER	choice prob. (H_9)	log P	max ER-ER	choice prob. (H_{10})	log P	choice prob.	log P	
1	1.55	-1.10	11.10	0	1.22	.8197	.069	-1.1612	3.99	.096	-1.0177	0	0	- ∞	.092	-1.0362	3
2	3.45	-1.15	9.20	.05	1.06	.9434	.079	-1.1024	4.15	.133	- .8761	.16	.048	-1.3188	.040	-1.3979	0
3	5.30	-1.20	7.35	.10	.90	1.111	.093	-1.0315	4.31	.138	- .8601	.32	.096	-1.0177	.046	-1.3372	2
4	7.15	-1.25	5.50	.15	.74	1.3514	.113	- .9469	4.47	.143	- .8447	.48	.144	- .8416	.050	-1.3010	2
5	8.95	-1.30	3.70	.	.59	1.6949	.142	- .8477	4.62	.148	- .8297	.64	.189	- .7235	.040	-1.3979	1
6	10.80	-1.35	1.85	.29	.43	2.3256	.195	- .7100	4.78	.153	- .8153	.79	.237	- .6253	.062	-1.2076	1
7	12.65	-1.40	0	.30	.27	3.7037	.310	- .5086	4.94	.156	- .8013	.95	.284	- .5467	.670	- .1739	15
prob.	1/9	8/9	1/9	8/9													
mats	12.65	-1.10			1.22												
Σ					5.21	11.9498			31.26			3.33					24
log L_{H_i}							-16.6271			-20.1272			-∞			-13.5990	log L_{H_0} = 24•(-.8447) = -20.2728

log $LR_{1/9}$ = 6.5282   log $LR_{0/1}$ = 6.6738

log $LR_{1/8}$ = 3.0281   log $LR_{0/9}$ = 0.1456

log $LR_{8/9}$ = 3.5001   log $LR_{0/8}$ = 3.6457

Table 22

likelihood ratio between		more favored hypothesis		
		H_1: diluted EV maximization	H_8: reciprocal regret	H_9: sum-diff. regret
less favored hypothesis	H_8: reciprocal regret	$1.067 * 10^3$		
	H_9: sum-diff. regret	$3.375 * 10^6$	$3.163 * 10^3$	
	H_0: random choice	$4.719 * 10^6$	$4.423 * 10^3$	1.111

have been provided beforehand.

Pitz, 1968 found another way of handling the problem of data probabilities of 0 and 1, in another context, but also with data originally not observed with a likelihood ratio analysis in mind. He tested a (null-) hypothesis H_o of equal probability of two kinds of observations ($p = 0.5$) against the rather unspecific hypothesis H_1 of $p > 0.5$. The data showed that 32 out of 48 Ss gave responses in accordance with H_1. The likelihood ratio for these data would have been

$$L = \frac{.5^{48}}{p_1^{32} (1-p_1)^{16}}$$

From this equation Pitz determined the value of p_1 for which the data would be equivocal, i.e., for which L would be one: $.5^{48} = p_1^{32} (1-p_1)^{16} \Rightarrow p_1 \approx .8$. (That means: if H_1 meant $p > .8$, the data would actually favor H_o rather than H_1.) Pitz's suggestion is to consider H_1 as a distribution $g(p)$ over p rather than a constant p_1, such that the likelihood ratio is

$$L = \frac{.5^{48}}{\int_{.5}^{1.0} p^{32} (1-p)^{16} g(p)\, dp},$$

and he proposes several possible distributions $g(p)$, such as a uniform (rectangular) distribution over [.5, 1.0], a triangular distribution with $g(p) = 0$ for $p \leq .5$, and a kind of beta distribution with a rather high mean. Such an analysis could be done with the Seghers, Fryback & Goodman data, too.

Conclusion

Now that we have seen that we can figure likelihood ratios between various competing hypotheses on given data sets which were not even made for it, <u>what do we do now</u>?

For a complete Bayesian data analysis, we would multiply our computed likelihood ratios to some prior odds for the respective hypotheses. These prior odds may be more or less public, or may be our very personal belief states. Methods to elicit and assess such prior distributions have been introducted and discussed elsewhere (e.g., Winkler, 1967, Staël von Holstein 1970).

For a complete Bayesian analysis, we would consider the possible consequences of our decisions between competing hypotheses, in terms of utilities assessed to the various combinations of our decisions among hypotheses with the possible "true" states of the world, and use these utilities in connection with our prior odds to determine cutoffs for the likelihood ratios where to decide in favor of which hypothesis or model. There are various techniques available now for the assessment of utilities to outcomes, even if these outcomes are characterized by several revelant attributes. These techniques have been summarized recently by Fischer (1972).

As we have seen in the few examples given in this paper, likelihood ratios grow rather rapidly with larger amounts of data. Even very biassed prior odds would be brought very soon into the correct range by multiplication to these large likelihood ratios. This indicates that Bayesian analyses might get along with much smaller sample sizes than traditional statistical data analyses

with their diffuse alternative hypotheses. How much precisely can be economized on the sample size, will depend in each case on the cutoff determined by prior odds and costs and payoffs (utilities) involved, as indicated by a proper decision analysis (see, e.g., Raiffa, 1969). That a careful formulation of competing hypotheses alone can result in considerable savings on expected sample size, has been shown by Wald (1947) already.

References

Atkinson, R. C., Bower, G. H., & Crothers, E. J. An introduction to mathematical learning theory. New York: Wiley, 1965.

Bush, R. R. Estimation and evaluation. Handbook of Mathematical Psychology, Ch. 8, Vol. 1. Wiley, New York, 1963.

Cunningham, D. R., & Briepohl, A. M. Empirical Bayesian learning. IEEE Transactions on Systems, Man, and Cybernetics, SMC-1 (1), 19-23, 1971.

Edwards, W., Lindman, H., & Savage, L. F. Bayesian statistical inference for psychological research. Psychol. Rev., 1963, 70, 193-242.

Fu, K. S. Sequential methods in pattern recognition and machine learning. Academic Press, New York & London, 1968.

Gettys, C. F., & Willke, T. A. The application of Bayes's Theorem when the true data state is unknown. Organizational Behavior and Human Performance, 1969, 4, 125-141.

Kuhn, T. S. The structure of scientific revolutions. University of Chicago Press, 1962.

Lindley, D. V. A statistical paradox. Biometrika, 1957, 44, 187-192.

McGee, V. E. Principles of statistics: Traditional and Bayesian. Appleton-Century-Crofts, New York, 1971.

Novick, M. R., & Grizzle, J. E. A Bayesian approach to the analysis of data from clinical trials. J. Amer. Stat. Ass., 1965, 60, 81-95.

Pitz, G. F. An example of Bayesian hypothesis testing: the perception of rotary motion in depth. Psychol. Bull., 1968, 70, 252-255.

Ramsey, F. L. A Bayesian approach to bioassay. Biometrics, 1972, 28, 841-858.

Restle, F. & Greeno, J. Introduction to Mathematical Psychology. Addison-Wesley, Reading, Mass. 1970.

Staël von Holstein, C-A. S. Assessment and evaluation of subjective probability distribution. The Economic Research Institute at the Stockholm School of Economics, Stockholm. 1970.

Sutcliffe, J. P. A general method of analysis of frequency data for multiple classification designs. Psychol. Bull., 1957, 54, 134-137.

Wald, A. Sequential Analysis. New York: Wiley, 1947.

Weisberg, H. I. Bayesian comparison of two ordered multinominal populations. Biometrics, 1972, 28, 859-867.

Willis, R. E. A Bayesian framework for the reporting of experimental results. Decision Sciences, 1972, 3 (4), 1-18.

Winkler, R. L. The assessment of prior distributions in Bayesian analysis. J. Amer. Stat. Ass., 1967, 62, 776-800.

Winkler, R. Introduction to Bayesian Inference and Decision. Holt-Reinhart & Winston, New York. 1972.